

# 科创 8

## 一、蒙特卡洛模拟、接受度及其计算

蒙特卡洛模拟 (Monte Carlo, MC) 就是先在计算机里产生大量“已知真值”的宇宙线粒子, 再让这些粒子按物理过程穿过 AMS 探测器。模拟中会包括粒子入射、经过探测器材料、产生信号、被重建以及最后通过筛选这些步骤。

真实飞行数据只告诉我们探测器最后记录到了什么。它不能直接告诉我们粒子原本的刚度是多少、入射方向是多少, 也不能告诉我们还有多少粒子本来来了但没有被接收到。这些信息只能从 MC 的真值记录里去估计。所以在通量分析里, MC 的作用是给飞行数据补上探测器效应这部分信息, 也就是回答“探测器把真实粒子变成测量结果时改动了多少”。

具体到本题, MC 主要有几件用处: 计算有效接受度, 估计重建和筛选效率, 建立刚度响应矩阵, 另外还要帮助估计一部分碎裂本底。它本身也要接受检查, 比如和飞行数据、束流测试结果比较, 看材料描述、核相互作用截面和分辨率模型是否合理。如果 data 和 MC 有差别, 还要把这个差别作为效率修正或系统误差处理。

探测器接受度可以理解为 AMS 对某类粒子的接收能力。它包含几何面积、允许的入射方向, 以及粒子经过重建和筛选后还能留下来的概率。通量公式可写成

$$\Phi_i = \frac{N_i}{A_i \epsilon_i T_i \Delta R_i}$$

其中  $N_i$  是第  $i$  个刚度 bin 中经过背景扣除和刚度迁移修正后的事例数,  $A_i$  是有效接受度,  $\epsilon_i$  是触发效率或其他没有并入  $A_i$  的效率因子,  $T_i$  是曝光时间,  $\Delta R_i$  是 bin 宽。这里要注意, 哪些效率放进  $A_i$ , 哪些单独写成  $\epsilon_i$ , 需要在分析中说清楚, 否则容易把同一个效率算两次。

接受度一般用 MC 来算。做法是在 AMS 上方放一个足够大的产生平面, 保证向下穿过 AMS 的粒子都会先经过这个平面。PPT 中这个平面是边长  $L = 3.9\text{ m}$  的正方形, 高度约  $z = +1.95\text{ m}$ , 对应几何因子约  $47.784\text{ m}^2\text{ sr}$ 。接着在平面上各向同性地产生粒子, 让它们经过完整模拟、重建和与数据相同的筛选。如果某个刚度 bin 中产生了  $N_{\text{gen},i}$  个 MC 事例, 最后留下  $N_{\text{sel},i}$  个, 则

$$A_i \simeq A_{\text{plane}} \frac{N_{\text{sel},i}}{N_{\text{gen},i}} C_i$$

这里  $A_{\text{plane}}$  是产生平面的几何因子,  $C_i$  表示 data/MC 效率差异、核相互作用损失或存活概率等修正。几何因子只管理想几何接收范围; 有效接受度还把筛选、重建和材料效应算了进去, 所以更接近实际测量中 AMS 对目标粒子的接收能力。

## 二、为什么要加事例筛选条件？主要背景是什么？

事例筛选需要在信号效率和样本纯度之间取一个比较合适的平衡。如果完全不筛选，留下的事例很多，但里面会混入非目标核子、碎裂事件和错误重建事件；如果筛选太严，本底少了，真正的信号也会被删掉很多。最后要让通量误差尽量小，而不是让某一个 cut 看起来最干净。

筛选条件类型	物理目的
<b>Inner Tracker hit 数要求</b>	L2 到 L8 中要有足够多击中点，例如至少 5 个 hit，并且最好分布在不同层段。这样轨迹拟合才稳定，也能减少偶然击中或错误组合造成的假轨迹。
<b>L1/L9 的 <math>x, y</math> hit 要求</b>	粒子要穿过指定几何，并在 $x, y$ 两个方向都有位置测量。L1 可以检查上方电荷和轨迹外推；FullSpan 分析中加入 L9 后轨迹臂长更长，高刚度下的测量会更有利。
<b>轨迹 <math>\chi^2</math> 质量要求</b>	Inner Tracker 拟合质量、L1 与 Inner Tracker 的联合拟合质量都要足够好。拟合很差的事例往往对应强散射、碎裂、错误匹配或重建质量差。
<b>电荷一致性筛选</b>	Inner Tracker、L1、UTOF 等子探测器测得的电荷要接近目标 $Z$ ，例如要求 $ Q - Z $ 小于一定范围。多个电荷测量相互一致时，相邻核子的污染会明显降低。
<b>刚度和地磁相关要求</b>	粒子刚度要高于相应地磁截止影响区间。这样可以避免把明显受地磁场影响的低刚度粒子直接拿来代表原始宇宙线能谱。

即使做完这些筛选，样本里还是会有背景。课件里重点讲的是碎裂背景。第一类是 Below L1 背景：较高  $Z$  的原子核在 L1 以下、Inner Tracker 以上的 TRD 和 Upper TOF 等材料中碎裂，变成目标核子，然后通过后面的筛选。这部分还可以利用 L1 电荷响应模板去拟合数据。因为 L1 和 L2 的硬件结构比较接近，而 L2 到 L8 之间物质量较少，所以可以用较纯的样本来建立模板。第二类是 TOI / Above L1 背景：较高  $Z$  原子核在 Tracker L1 或更上方的支撑结构中已经碎裂成目标核。L1 上方没有更多电荷测量层，飞行数据很难直接把它分出来，所以通常主要靠 MC 估计，再用飞行数据修正 MC 中碎裂分支比的 data/MC 差异。除此之外，有限电荷分辨率造成的残余误认也要检查，不过多探测器电荷一致性筛选通常能把这部分压低。

因此背景修正不能简单写成减去一个固定数。它和碎裂位置、母核种类、测量刚度、探测器分辨率以及 data/MC 差异都有关系。TOI 背景尤其要按测量刚度处理，因为碎裂和刚度分辨率会让一个真实 10 GV 的粒子最后被重建到 8 GV 或其他刚度 bin。

### 三、刚度迁移修正 unfolding 是什么？

Tracker 测刚度有有限分辨率，所以探测器重建出来的刚度不一定等于粒子的真实刚度。真实刚度在某个 bin 的粒子，可能被重建到相邻 bin；如果分辨率有非高斯尾部，也可能跑到更远的 bin。宇宙线能谱随刚度升高下降很快，低刚度区事件多，高刚度区事件少，因此哪怕只有一小部分粒子迁移，也可能改变测得的谱形。

Unfolding 要做的事，是从测量分布里去掉探测器分辨率造成的模糊，尽量恢复真实刚度分布。它不是把每个事件逐个改回真实刚度，因为单个事件的真实刚度无法唯一确定；它处理的是 bin 与 bin 之间的统计关系。这个关系可以写成响应矩阵形式：

$$n_i^{\text{obs}} = \sum_j R_{ij} N_j^{\text{true}} + b_i$$

其中  $R_{ij}$  表示真实在第  $j$  个 bin 的粒子被重建到第  $i$  个 bin 的概率，通常由 MC 给出； $b_i$  是背景。扣除背景后，再利用响应矩阵反推出  $N_j^{\text{true}}$ ，这就是 unfolding 的基本想法。从数学上看，它类似反卷积：真实分布被探测器分辨率卷积成测量分布，unfolding 再在统计意义上把这个分辨率效应扣回来。

所以最后放进通量公式的  $N_i$ ，不应该是原始数出来的事例数，而应该是背景扣除和刚度迁移修正后的事例数。这样算出的通量才更接近进入 AMS 探测器顶部时的真实宇宙线通量。